Khamis, M., Hasholzner, R., Bulling, A. and Alt, F. (2017) GTmoPass: Two-factor Authentication on Public Displays Using Gaze-touch Passwords and Personal Mobile Devices. In: 6th ACM International Symposium on Pervasive Displays, Lugano, Switzerland, 7-9 Jun 2017, p. 8. ISBN 9781450350457.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

© The Authors 2017. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in Proceedings of the 6th ACM International Symposium on Pervasive Displays, Lugano, Switzerland, 7-9 Jun 2017, p. 8. ISBN 9781450350457 https://doi.org/10.1145/3078810.3078815.

http://eprints.gla.ac.uk/170220/

Deposited on: 5 October 2018

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk

GTmoPass: Two-factor Authentication on Public Displays Using Gaze-Touch passwords and Personal Mobile Devices

Mohamed Khamis[1], **Regina Hasholzner**[1], **Andreas Bulling**[2], **Florian Alt**[1]
[1]Ubiquitous Interactive Systems Group, LMU Munich, Germany
[2]Max Planck Institute for Informatics, Saarland Informatics Campus, Germany
{mohamed.khamis, florian.alt}@ifi.lmu.de, bulling@mpi-inf.mpg.de

ABSTRACT
As public displays continue to deliver increasingly private and personalized content, there is a need to ensure that only the legitimate users can access private information in sensitive contexts. While public displays can adopt similar authentication concepts like those used on public terminals (e.g., ATMs), authentication in public is subject to a number of risks. Namely, adversaries can uncover a user's password through (1) shoulder surfing, (2) thermal attacks, or (3) smudge attacks. To address this problem we propose GTmoPass, an authentication architecture that enables Multi-factor user authentication on public displays. The first factor is a knowledge-factor: we employ a shoulder-surfing resilient multimodal scheme that combines gaze and touch input for password entry. The second factor is a possession-factor: users utilize their personal mobile devices, on which they enter the password. Credentials are securely transmitted to a server via Bluetooth beacons. We describe the implementation of GTmoPass and report on an evaluation of its usability and security, which shows that although authentication using GTmoPass is slightly slower than traditional methods, it protects against the three aforementioned threats.

Author Keywords
Multi-factor Authentication, Pervasive Displays, Eye Gestures

ACM Classification Keywords
K.6.5 Computing Milieux: Security and Protection: Authentication

INTRODUCTION
Public displays deliver various kinds of tangible benefits and are now deployed in train stations, airports, and streets. Meanwhile, there is an increasing demand for displays to offer personalized, context-specific content [7, 21, 23].

There are many cases in which users need to securely authenticate at public displays. For example, while a group of tourists examine places to visit on a large public display, the system could allow users to buy tickets for museums, buses, etc. One or more users could then authenticate in parallel by entering their passwords on their mobile devices. Using the mobile device's MAC address and the provided password, the system validates the user's credentials and charges the correct account for the ticket fee. While there are scenarios where it might be acceptable to continue the purchase and interaction on the mobile device, in many cases it is favorable to keep the user at the display to resume the primary task. In the aforementioned example, tourists could then be shown further suggestions for activities at the given place.

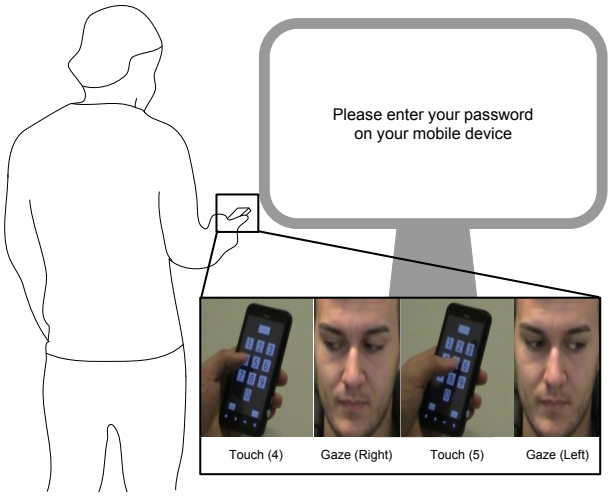

Figure 1. We propose a Multifactor authentication architecture for public displays. Users authenticate by entering a shoulder-surfing resilient GazeTouch password (knowledge factor) on their personal mobile device (possession factor). Users simply need to launch the GazeTouchPass app [19] that we modified so that it would securely communicate the password to an authentication server, whose URL is received by the phone through an Eddystone-URL broadcasted using a BLE Beacon.

When exchanging sensitive data with a public display (e.g., login credentials), users are prone to several types of threats. Namely, adversaries can uncover a user's password in public space through: (1) *Shoulder surfing attacks*: observing the user while authenticating [14], (2) *Thermal attacks*: exploiting the heat traces resulting from the user's interaction with the interface [1, 24], and (3) *Smudge attacks*: exploiting the oily residues left after authenticating on touchscreens [3]. While the latter two risks were demonstrated to be feasible, shoulder surfing was shown to occur in daily contexts [14].

In this work we introduce GTmoPass, an authentication architecture that enables multi-factor user authentication on public

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

PerDis '17, June 07-09, 2017, Lugano, Switzerland

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ISBN 978-1-4503-5045-7/17/06…$15.00
DOI: http://dx.doi.org/10.1145/3078810.3078815

displays. GTmoPass uses a shoulder-surfing resilient multimodal scheme that combines **G**aze and **T**ouch for password entry as a knowledge factor. Additionally it uses personal **mo**bile devices as a possession factor (see Figure 1). After entering the password on the mobile device, the password is then securely transferred to an authentication server whose URL is communicated to the mobile device via Bluetooth beacons. The use of BLE beacons alleviates the need to manually enter URLs, or scan QR-codes. This means that when interacting with public display that employs GTmoPass for the first time, users do not have to do anything other than launching the app and entering the password.

The results of our evaluation show that users authenticate relatively fast (2.8 – 4.9 seconds), and that the authentication process is resilient to the aforementioned threats. Even if the password is leaked, the architecture of GTmoPass requires the adversary to additionally acquire the user's mobile device.

BACKGROUND AND RELATED WORK

Authentication Factors

Researchers and practitioners have developed different ways for users to authenticate in order to be granted access to Private or sensitive information. Three of the most popular authentication forms are (1) knowledge-based authentication, (2) possession-based authentication, and (3) inherence-based authentication (also known as biometric authentication).

Knowledge Factor

Knowledge-based authentication schemes rely on "something the user knows". It is perhaps the most commonly used factor [2]. Examples are passwords, PINs, and graphical passwords. Researchers also developed ways to authenticate using eye movements [8, 12, 13, 16], mid-air gestures [25], and by recalling photographs [26]. Knowledge-based schemes allow changing passwords, and can be integrated into any system that accepts any kind of user input. On the other hand, knowledge-based passwords can be guessed by illegitimate users. Attackers can find the password by observing users during authentication [14]. Smudge attacks are possible when passwords are entered through touchscreens [3]. Graphical passwords such as Android lock patterns are particularly vulnerable to smudge attacks [29, 40]. Furthermore, many knowledge-based schemes are also vulnerable to thermal attacks, where heat traces resulting from the user's interactions with the interface are exploited to find the password [1].

Possession Factor

The possession factor relies on "something the user possesses". Physical keys and scannable personal IDs are examples of possession-based authentication. Researchers experimented with identifying users on public displays through the MAC address of their smartphone [28]. Davies et al. [7] exploit the user's mobile device to identify the list of applications the user wants to interact with on a display. Others approaches include using bluetooth devices [6] and wearable shutter glasses [32]. While this type of schemes does not require users to remember passwords, it requires keeping possession of the access token. A drawback of using this factor alone, is that if an attacker gets hold of the access token, the attacker can impersonate the user and gain unauthorized access.

Inherence Factor

The inherence factor relies on biometric data, such as fingerprints, user behavior (e.g., behavior on a touchscreen [10] or eye-gaze behavior [33]) and face detection [15]. While biometric authentication can be easy and fast to use, it is accompanied with a number of problems. Biometric passwords cannot be changed; once a user's biometric data (e.g., fingerprint or iris scan) is leaked, there is no way the user can invalidate the leaked data. Face recognition can be bypassed using pictures of the legitimate user [22], and users leave fingerprints everywhere as they interact with surrounding objects. Furthermore, users are oftentimes concerned about disclosing their biometric data to third-party companies [27], especially after it was found that biometric data can be stolen remotely [34, 42].

Multifactor Authentication

Multifactor authentication refers to the use of two or more of the aforementioned factors for improved security. This approach is employed by ATM machines; users have to *know* a PIN, and have to *possess* an ATM card. The approach has also been adopted by multiple Internet services, such as Google, Facebook, Microsoft, and Dropbox; users have to *know* their username and password, and have to *possess* a previously identified mobile device on which they receive an additional one-time password, or confirm their log-in attempt.

Researchers developed systems where users authenticate at ATMs by entering PINs on their mobile phones [4, 30, 31]. De Luca and Frauendienst introduced PocketPIN, where users can enter credit card data on their phones before being securely transmitted to public terminals [9].

The advantage of GTmoPass is that it employs multimodal authentication as a knowledge factor, and personal mobile devices as a possession factor. Multimodal authentication was shown to be highly resilient to shoulder surfing [19]. Furthermore, thermal and smudge attacks normally require the attacker to inspect the interface after the user had left [1, 3]. Our architecture complicates these attacks by relying on the user's mobile device for input. This means that an attacker can only perform these attacks by stealing the mobile device fast enough before the heat or smudge traces can no longer be traced. And even by doing so, the attacker would not be able to identify the gaze-input.

Protecting Privacy on Public Displays

In addition to the aforementioned works by Davies et al. [6, 7] and Schaub et al. [28], other works exploited proxemics for showing content on public displays. For example, Vogel and Balakrishnan show private content on public displays only when the user is very close to the display [36]. Brudy et al. proposed some concepts to hide private data on public displays by partially hiding the private data from the observer's view estimated by a Kinect [5].

GTMOPASS

GTmoPass is an authentication architecture that enables secure multifactor user authentication on public displays. In the

following we describe the concept and implementation of GTmoPass, and which threat models it is optimized against.

Concept

GTmoPass relies on two factors (1) a knowledge factor: we use a multimodal authentication scheme that combines gaze and touch input, and (2) a possession factor: users enter the multimodal passwords on their personal mobile devices.

The Knowledge Factor

For the knowledge factor in GTmoPass, we employ a modified version of GazeTouchPass [19], a state-of-the art authentication scheme that is highly resilient to shoulder surfing attacks. GazeTouchPass is a multimodal scheme that employs combinations of touch-based PINs (0-9) and eye movements (gaze gestures to the left and to the right). This means that it uses a theoretical password space of 12^n (10 digits + 2 gaze gestures) where n is the length of the password. A password could be: Gaze (left), Touch (2), Gaze (right), Touch (1).

The strength of GazeTouchPass lies in its use of two input modalities. This adds complexity to shoulder surfing attacks, as it requires the attacker to observe (1) the user's input on the touchscreen, and (2) the eye movements of the user.

The Possession Factor

For the possession factor in GTmoPass, we rely on the user's personal mobile device. The multimodal passwords are entered on the mobile device; touch input is detected through the touchscreen, and gaze input is detected through the front-facing camera. The mobile device then communicates securely with an authentication server, that validates the password and signals the display to show the private information.

Implementation

GTmoPass consists of two main components: (1) the authentication server, and (2) the user's mobile device (client).

Authentication Server

Implemented in NodeJS, the authentication server is configured to receive HTTP requests. The server runs on a computer (e.g., in our setup we used a laptop) that is connected to a WiFi router. The IP address of the server is broadcasted using a BEEKS BLE beacon[1] in Google's Eddystone protocol[2]. The IP is broadcasted in Eddystone-URL data packets at a 10 Hz rate (i.e., it broadcasts once every 100 ms), with a transmission power of 0 dBm ($\approx$ 50 meters).

The Client

GazeTouchPass [19] was implemented as an Android application. It uses the OpenCV library[3] and the Viola-Jones classifier [35] to detect the user's face and eyes. Afterwards, in a manner similar to previous work [41, 43], the gaze direction is estimated depending on the distance between each eye and the center of the user's face.

We further extended GazeTouchPass to communicate with the authentication server. As soon as the modified app launches, it scans for beacons and connects to the nearest one that broadcasts Eddystone URLs. The URL is then saved in the local memory. Whenever the user provides four inputs, the system hashes the input using SHA 256 and sends an HTTP request to the server (see Figure 2). Similar to previous work [28], the mobile device is uniquely identified through its MAC address.

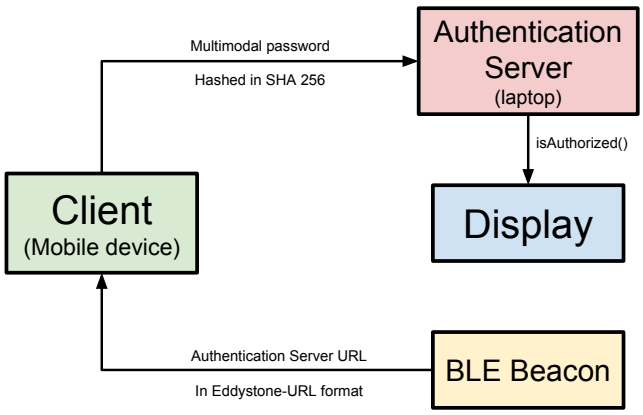

Figure 2. As a user approaches the display to be authenticated at, the user would take out his/her personal mobile device and launch the modified GazeTouchPass app. The app receives the Eddystone-URL broadcasted by the beacon and presents the log in screen. The user then enters the multimodal password. The password is hashed using SHA 256 and then securely transferred to the authentication server. The server validates the log in credentials and signals the display that the credentials were correct/incorrect.

EVALUATION

We previously evaluated the usability and observation resistance of GazeTouchPass, to find that although it requires 1.6 more seconds that PINs [39], it is significantly more secure against shoulder surfing compared to traditional PINs [19]. We also found that the structure of a multimodal password has an influence on its security. Namely, passwords that contain several switches from one modality to another, are more difficult to observe compared to those that have less switches. For example, a password such as "left-2-right-1" has three modality switches, and is hence harder to observe compared to "left-right-2-1", which has only one modality switch. The reason is that attackers would have to switch attention between the user's fingers and the user's eyes more often in the case of passwords with more modality switches.

Another interesting insight from our previous evaluation of GazeTouchPass, is that multimodal passwords that start or end with gaze input were perceived by participants to be more difficult to observe. While our previous study was designed to focus on the effect of the number of switches in input modalities on the observability of the password, in the currently presented work we focus on the influence of the *position* of the gaze input in the password on the usability and security of the GazeTouch password.

Usability Study

The goals of this study are to: (1) collect feedback about the use of GazeTouchPass in the proposed setup, and (2) understand the influence of the gaze-input's *position* on usability.

[1] http://bluvision.com/beeks/
[2] https://developers.google.com/beacons/eddystone-eid
[3] http://opencv.org/platforms/android.html

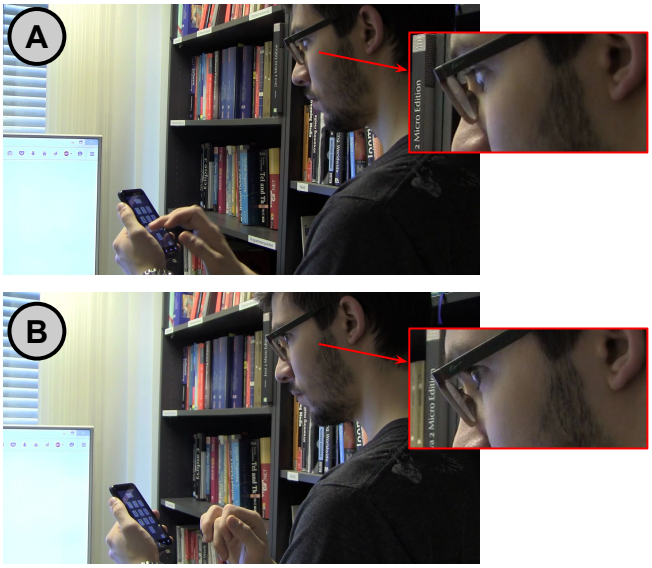

Figure 3. Participants were recorded during the usability study as they enter passwords using an HD video camera. The recorded videos were used in the subsequent security study to simulate shoulder surfing attacks. Figure A shows a user entering a touch input *Touch(7)*, while figure B shows a user performing a gaze gesture *Gaze(Left)*.

Apparatus

We used a 48 inch Samsung TV (1920×1080 px) as a display. We connected a laptop computer running a NodeJS server that accepts HTTP requests. The server validates the received passwords and updates the display's content accordingly. We recorded participants as they enter passwords from the side (see Figure 3). An HD camera was positioned such that it is close enough to show the touchscreen, and also the user's eyes. These videos were recorded to be used in the subsequent security study, to simulate shoulder surfing attacks.

Design

Since we wanted to investigate the influence of the position of the gaze input in the password, we experimented with four conditions: (1) passwords that start with gaze input (*GazeStart*), (2) passwords that end with gaze input (*GazeEnd*), (3) passwords that start and end with gaze input (*GazeStartEnd*), and (4) passwords with gaze input in the middle (*GazeMiddle*).

The study was designed as a repeated measures experiment. Each participant performed 16 authentications (4 passwords × 4 conditions) using randomly generated passwords. In case of conditions GazeStart and GazeEnd, participants entered two passwords with one gaze input at the start/end of the password, while the other two passwords had two gaze inputs at the start/end of the password. Table 1 shows sample passwords.

Participants

We invited 16 participants to our lab (6 females), recruited through mailing lists and social networks. Participants were awarded with online shop vouchers or participation points. All participants had normal or corrected-to-normal vision.

Procedure

The experimenter first described the study and asked the participants to sign a consent form. She then handed the participant a mobile device with the modified version of GazeTouchPass installed, and explained how it works. Each participant was allowed to perform a training run per condition to get acquainted with the system. The trial attempts were not included in the analysis. At each authentication attempt, the experimenter read out the password to be entered according to a previously generated list that was randomized. The participant would then enter the password, and observe the feedback on the display, which indicated whether or not the correct password was detected. Afterwards participants were interviewed to learn about their feedback, ideas and concerns about GTmoPass.

Condition		Format				Example
GazeStart	startOne	Gaze	Touch	Touch	Touch	L123
	startTwo	Gaze	Gaze	Touch	Touch	RL34
GazeEnd	endOne	Touch	Touch	Touch	Gaze	943L
	endTwo	Touch	Touch	Gaze	Gaze	53RR
GazeMiddle		Touch	Gaze	Gaze	Touch	7LL3
GazeStartEnd		Gaze	Touch	Touch	Gaze	R82L

Table 1. Previous work reported that the position of the gaze input was perceived by participants to have an influence on the password's observability [19]. Hence we experimented with the above conditions to cover possible positions of the gaze input in the password.

Results

To measure the impact of the position of gaze input on the usability of the passwords, we measured the *entry time* and the *error count* while authenticating.

Entry time was measured starting from the first input until the last input was recognized. Figure 4 illustrates the time taken to enter a password at every condition. A repeated measures ANOVA (with Greenhouse-Geisser Correction due to violation of sphericity) showed a significant effect of the gaze input's position on the time it takes to enter a password ($F_{1.8,27.5} = 9.1$, $p < 0.05$). Post-hoc analysis with Bonferroni correction (α = 0.05 / 6 comparisons = 0.0083) showed significant differences in entry time between GazeStart ($M = 2863$ ms, $SD = 1525$ ms) and GazeMiddle ($M = 4959$ ms, $SD = 3141$ ms), between GazeEnd ($M = 3892$ ms, $SD = 3045$ ms) and GazeMiddle ($M = 4959$ ms, $SD = 3141$ ms), and between GazeStartEnd ($M = 3757$ ms, $SD = 3852$ ms) and GazeMiddle ($M = 4959$ ms, $SD = 3141$ ms). The other pairs were not significantly different ($p > 0.05$). This means that passwords with gaze in the middle are significantly slower than other cases.

Error count reflects the number of times the participant entered the password incorrectly. Errors could occur either due to entering the wrong gaze or touch input, or due to the system detecting an incorrect gaze input due to poor lighting conditions. Figure 5 shows the number of errors at each condition.

Qualitative Feedback collected at the end of the study through semi-structured interviews revealed positive feedback towards the system. Many participants reported that they liked the idea of detecting eye movements through the smartphone's camera, and would imagine using it to authenticate on ATMs instead of using cards and PINs Some participants suggested using the system to open security doors. One participant suggested using it for authentication on other digital devices.

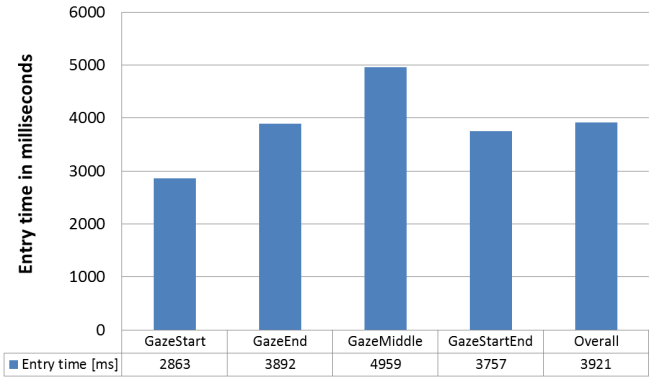

Figure 4. Mean authentication times.

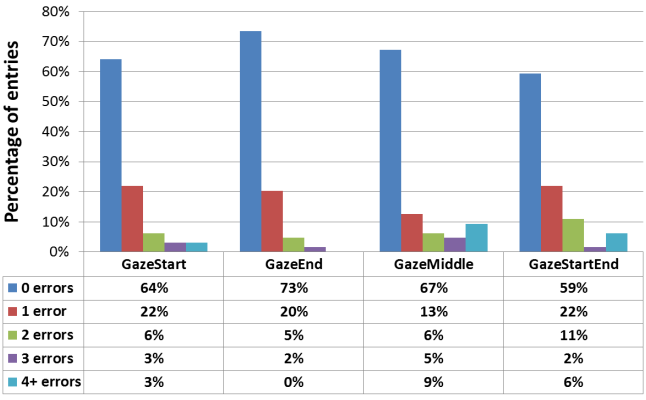

Figure 5. Number of attempts before a successful entry.

Security Study

Unlike existing systems, GTmoPass is resilient to to thermal and smudge attacks by design. Heat traces and oily residues can only uncover the *touch* part of the password, but not the gaze input. Therefore in this study we focus on focus on GTmoPass's resilience to observation attacks.

Although previous work evaluated the impact of the number of modality switches on the security of passwords [19]. Our aim in this study was to understand the influence of gaze input's position on the observation resistance of the password.

Using the video material produced in the first study, we conducted a second *observability study* that simulated a shoulder surfing attack against a user authenticating using GTmoPass. To do this, we invited 16 different participants and asked them to simulate shoulder surfers by watching the recorded videos, and trying to find the entered passwords (see Figure 6).

Threat Model

In our threat model, the user and the attacker are in public space. The attacker is familiar with GTmoPass and how it works. The attacker observes the user from an angle that allows seeing both the touch input on the mobile device, and the eye movements (see Figure 3). The distance between the attacker and the user is close enough to see the touchscreen, but far enough to reduce the effort of switching focus back and forth between the user's eyes and the device's touchscreen. After unveiling the password, the attacker tries to get hold of the device and authenticate at the display.

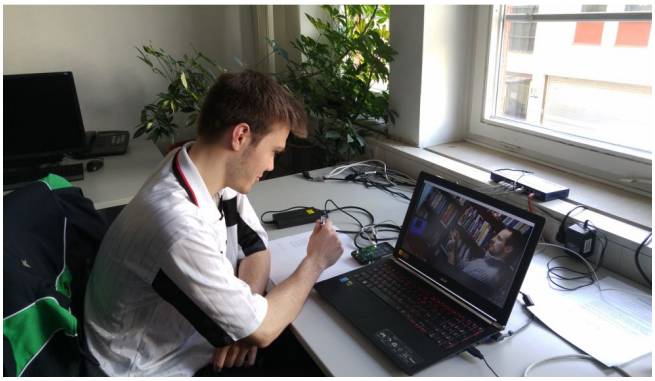

Figure 6. Participants of the security study watched HD videos of the usability study participants as they authenticated. The task was to find the correct password. Participants were explained how the system works, had a chance to try the application themselves, and were allowed to take notes while watching the videos.

Participants and Reward mechanism

We invited 16 participants (9 females) through mailing lists and social networks. None of them had participated in the usability study. All participants were awarded either an online shop voucher or participation points. In addition, they took part in a lottery for an additional 10 EUR online shop voucher, where the chance of winning increases as the participant successfully attacks more passwords. This was done to encourage participants to put a lot of effort in finding the passwords.

Design

This study also followed a repeated measures within subjects design. Each participant watched 4 videos of successful authentications from each condition (4 videos $\times$ 4 conditions = 16 videos in total), each of which is a recording from a different usability study participant.

Procedure and Apparatus

After arriving at our lab and filling out the consent form, the experimenter explained the concept and the reward mechanism. The videos were displayed and controlled by the experimenter on a computer (1920 px $\times$ 1080 px). Participants were given a pen and a paper and were allowed to take notes while watching the videos (see Figure 6). Since each video was attacked once, it was watched once and hence the duration of the attack depends on the length of the video. They were also allowed to try the application themselves. After watching the video once, they provided up to three guesses, but were not told whether their guesses were correct or not to avoid influencing the perceived difficulty. We concluded with a questionnaire to learn more about the perceived difficulty of attacks.

Results

To understand the impact of the position of gaze input on the observability of the passwords, we measured the *binary success rate* and the *Levenshtein distance*.

Success rate reflects how many passwords of each condition were successfully observed. On the other hand, the *Levenshtein distance* indicates how close the guess is to the original password [1, 19, 38]. Since our participants provided 3 guesses against each password they observed, only the guess

Mean values	GazeStart	GazeEnd	GazeMiddle	GazeStartEnd
Binary success rate (0%=all incorrect;100%=all correct)	19%	16%	13%	16%
Levenshtein distance (0=similar;4=completely different)	1.36	1.18	1.21	1.27
Perceived Difficulty (1=very easy;4=very hard)	4	3	3.5	4

Table 2. We measured the binary success rate (i.e., whether or not the attacker's guess is correct) and the Levenshtein distance (i.e., how similar the guesses are to the correct password). A low success rate and high Levenshtein distance indicate high observation resistance and hence higher security. The most secure passwords are the GazeMiddle, where only 13% of the passwords were observed successfully, and guesses were 55% similar to the correct ones. Participants reported the perceived difficulty of attacks before knowing whether their guesses were correct.

with the shortest distance (i.e., highest similarity) to the correct password was considered for analysis.

Table 2 shows the average success rate and Levenshtein distance for each condition. While participants succeeded the most when attacking GazeStart, their guesses had the longest average distance from the correct password. Guesses against GazeEnd were closest to the correct password (i.e., low distance) possibly because the attackers knew that each password consists of 4 inputs, and after seeing two or three touch-inputs they foresaw that the following inputs are gaze-based.

We also collected the *perceived difficulty* which participants indicated through a Likert scale (5-points;1=very easy;5=very hard). Table 2 shows that participants found all types difficult to observe, but GazeStart and GazeStartEnd were perceived to be more difficult compared to GazeEnd and GazeMiddle. These values are in-line with the Levenshtein distances.

DISCUSSION

The proposed architecture enables usable and secure authentication on public displays. Users simply approach the display with their unmodified personal mobile device that has our app installed. The app retrieves the IP of the server that is broadcasted by the BLE beacon, allowing the user to directly authenticate without entering URLs or scanning QR codes.

A usability study shows that the use of multimodal passwords in that setup is feasible, well perceived, and is only slightly slower than the less secure PINs (von Zezschwitz et al. report 1.5 seconds for PIN-entry [39]). While users make few errors, previous work has shown that users are willing to correct errors on public displays [20]. A security study showed that authentication is robust against observation attacks (only 13%-19% successful attacks in optimal conditions), which is more secure than PINs and several recently proposed systems. For example, attacks against EyePassShapes [8], EyePIN [12], GazeTouchPass [19] and XSide [11] had success rates of 42%, 55%, 19% and 9% – 38% respectively. Furthermore, the fact that gaze-input does not leave traces on the device makes GTmoPass secure against thermal and smudge attacks even if the attacker gets hold of the mobile device.

Trade-off between Usability and Security

While authentication using PINs is fast, it is known to be insecure and highly vulnerable to observation attacks [19, 37] and thermal attacks [1]. On the other hand authentication using multimodal passwords is more secure, but takes longer time compared to PINs. While even a slight increase in authentication time on mobile devices has a big impact considering that users unlock their mobile devices more than 50 times a day [18], we argue that authentication on public displays does not happen as often and hence a slight increase (between 1.3 and 2.5 seconds in our case) is not very significant.

In addition to the overall trade-off, we found that having gaze-input in the middle of the password (GazeMiddle) is the least likely to be successfully attacked, but also requires the longest time to enter. In general, it was found that providing consecutive gaze inputs results in longer authentication times. This was the case in GazeStartTwo and GazeEndTwo and GazeMiddle (see examples in Table 1). This is due to the time it takes to perform a gaze gesture, look to the front again, then perform another gaze gesture. On the other hand, guesses against GazeEnd are the closest to the actual password. We expect that after observing two or three touch inputs, participants foresaw that the following inputs could be gaze-based.

Perceived Difficulty of Shoulder Surfing

It is interesting that the perceived difficulty of attacks reported by participants was more in-line with the Levenshtein distances rather than with the binary success rate. The Levenshtein distance metric evaluates how similar a guess is to the actual password, which means that it also reflects how many times digits or gaze gestures were observed correctly. This means that unlike the binary success rate, participant's confidence in identifying particular inputs can be a valid indicator of low Levenshtein distances.

It is not surprising that the final success rate does not correlate with the perceived difficulty. In fact, previous work reported that attackers often underestimate the perceived difficulty of shoulder surfing. For example, in the work by George et al. [17], the perceived difficulty of performing shoulder surfing attacks changed drastically after trying to perform attacks.

CONCLUSION

In this work we showed that GTmoPass offers a secure authentication architecture for public displays. A usability and a security study showed that GTmoPass is usable and secure against shoulder surfing. We also discussed how thermal and smudge attacks are infeasible by design.

In the future, we want to evaluate more complicated threat models. For example, a combination of a thermal attack to uncover touch input and an observation attack to uncover gaze input, or multiple consecutive observations by insiders (e.g., family members or work colleagues). Another interesting threat model is the case of having two attackers: one observing the eyes, while the other observes the touchscreen. We also intend to conduct a field study to better understand how users perceive GTmoPass in the wild. A further direction for future work is to include a third inherence factor. This can be done by scanning the finger print or by face detection using the front-facing camera.

ACKNOWLEDGEMENT

This work was supported by a Google IoT Technology Research Award.

REFERENCES

1. Yomna Abdelrahman, Mohamed Khamis, Stefan Schneegass, and Florian Alt. 2017. Stay Cool! Understanding Thermal Attacks on Mobile-based User Authentication. In *Proceedings of the 35th Annual ACM Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3025453.3025461

2. M. D. H. Abdullah, A. H. Abdullah, N. Ithnin, and H. K. Mammi. 2008. Towards Identifying Usability and Security Features of Graphical Password in Knowledge Based Authentication Technique. In *2008 Second Asia International Conference on Modelling & Simulation (AMS)*. 396–403. DOI: http://dx.doi.org/10.1109/AMS.2008.136

3. Adam J. Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M. Smith. 2010. Smudge Attacks on Smartphone Touch Screens. In *Proceedings of the 4th USENIX Conference on Offensive Technologies (WOOT'10)*. USENIX Association, Berkeley, CA, USA, 1–7. http://dl.acm.org/citation.cfm?id=1925004.1925009

4. Andrea Bianchi. 2011. Authentication on Public Terminals with Private Devices. In *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '11)*. ACM, New York, NY, USA, 429–430. DOI: http://dx.doi.org/10.1145/1935701.1935815

5. Frederik Brudy, David Ledo, Saul Greenberg, and Andreas Butz. 2014. Is Anyone Looking? Mitigating Shoulder Surfing on Public Displays Through Awareness and Protection. In *Proceedings of The International Symposium on Pervasive Displays (PerDis '14)*. ACM, New York, NY, USA, Article 1, 6 pages. DOI: http://dx.doi.org/10.1145/2611009.2611028

6. Nigel Davies, Adrian Friday, Peter Newman, Sarah Rutlidge, and Oliver Storz. 2009. Using Bluetooth Device Names to Support Interaction in Smart Environments. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*. ACM, New York, NY, USA, 151–164. DOI: http://dx.doi.org/10.1145/1555816.1555832

7. Nigel Davies, Marc Langheinrich, Sarah Clinch, Ivan Elhart, Adrian Friday, Thomas Kubitza, and Bholanathsingh Surajbali. 2014. Personalisation and Privacy in Future Pervasive Display Networks. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2357–2366. DOI: http://dx.doi.org/10.1145/2556288.2557287

8. Alexander De Luca, Martin Denzel, and Heinrich Hussmann. 2009. Look into My Eyes!: Can You Guess My Password?. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. ACM, New York, NY, USA, Article 7, 12 pages. DOI: http://dx.doi.org/10.1145/1572532.1572542

9. Alexander De Luca and Bernhard Frauendienst. 2008. A Privacy-respectful Input Method for Public Terminals. In *Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges (NordiCHI '08)*. ACM, New York, NY, USA, 455–458. DOI:http://dx.doi.org/10.1145/1463160.1463218

10. Alexander De Luca, Alina Hang, Frederik Brudy, Christian Lindner, and Heinrich Hussmann. 2012. Touch Me Once and I Know It's You!: Implicit Authentication Based on Touch Screen Patterns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 987–996. DOI:http://dx.doi.org/10.1145/2207676.2208544

11. Alexander De Luca, Marian Harbach, Emanuel von Zezschwitz, Max-Emanuel Maurer, Bernhard Ewald Slawik, Heinrich Hussmann, and Matthew Smith. 2014. Now You See Me, Now You Don't: Protecting Smartphone Authentication from Shoulder Surfers. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2937–2946. DOI: http://dx.doi.org/10.1145/2556288.2557097

12. Alexander De Luca, Roman Weiss, and Heiko Drewes. 2007. Evaluation of Eye-gaze Interaction Methods for Security Enhanced PIN-entry. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces (OZCHI '07)*. ACM, New York, NY, USA, 199–202. DOI: http://dx.doi.org/10.1145/1324892.1324932

13. Alexander De Luca, Roman Weiss, Heinrich Hussmann, and Xueli An. 2008. Eyepass - Eye-stroke Authentication for Public Terminals. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08)*. ACM, New York, NY, USA, 3003–3008. DOI: http://dx.doi.org/10.1145/1358628.1358798

14. Malin Eiband, Mohamed Khamis, Emanuel von Zezschwitz, Heinrich Hussmann, and Florian Alt. 2017. Understanding Shoulder Surfing in the Wild: Stories from Users and Observers. In *Proceedings of the 35th Annual ACM Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA. DOI:http://dx.doi.org/10.1145/3025453.3025636

15. FaceLock. 2013. FaceLock. http://www.facelock.mobi/. (2013). accessed 10 January 2017.

16. Alain Forget, Sonia Chiasson, and Robert Biddle. 2010. Shoulder-surfing Resistance with Eye-gaze Entry in Cued-recall Graphical Passwords. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1107–1110. DOI: http://dx.doi.org/10.1145/1753326.1753491

17. Ceenu Goerge, Mohamed Khamis, Emanuel von Zezschwitz, Marinus Burger, Henri Schmidt, Florian Alt, and Heinrich Hussmann. 2017. Seamless and Secure VR: Adapting and Evaluating Established Authentication Systems for Virtual Reality. In *Proceedings of the*

Network and Distributed System Security Symposium (USEC '17). NDSS. DOI: http://dx.doi.org/10.14722/usec.2017.23028

18. Marian Harbach, Alexander De Luca, and Serge Egelman. 2016. The Anatomy of Smartphone Unlocking: A Field Study of Android Lock Screens. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4806–4817. DOI: http://dx.doi.org/10.1145/2858036.2858267
19. Mohamed Khamis, Florian Alt, Mariam Hassib, Emanuel von Zezschwitz, Regina Hasholzner, and Andreas Bulling. 2016a. GazeTouchPass: Multimodal Authentication Using Gaze and Touch on Mobile Devices. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 2156–2164. DOI: http://dx.doi.org/10.1145/2851581.2892314
20. Mohamed Khamis, Ludwig Trotter, Markus Tessmann, Christina Dannhart, Andreas Bulling, and Florian Alt. 2016b. EyeVote in the Wild: Do Users Bother Correcting System Errors on Public Displays?. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16)*. ACM, New York, NY, USA, 57–62. DOI:http://dx.doi.org/10.1145/3012709.3012743
21. Thomas Kubitza, Sarah Clinch, Nigel Davies, and Marc Langheinrich. 2013. Using Mobile Devices to Personalize Pervasive Displays. *SIGMOBILE Mob. Comput. Commun. Rev.* 16, 4 (Feb. 2013), 26–27. DOI: http://dx.doi.org/10.1145/2436196.2436211
22. Yan Li, Yingjiu Li, Qiang Yan, Hancong Kong, and Robert H. Deng. 2015. Seeing Your Face Is Not Enough: An Inertial Sensor-Based Liveness Detection for Face Authentication. In *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security (CCS '15)*. ACM, New York, NY, USA, 1558–1569. DOI: http://dx.doi.org/10.1145/2810103.2813612
23. Nemanja Memarovic. 2015. Public Photos, Private Concerns: Uncovering Privacy Concerns of User Generated Content Created Through Networked Public Displays. In *Proceedings of the 4th International Symposium on Pervasive Displays (PerDis '15)*. ACM, New York, NY, USA, 171–177. DOI: http://dx.doi.org/10.1145/2757710.2757739
24. Keaton Mowery, Sarah Meiklejohn, and Stefan Savage. 2011. Heat of the Moment: Characterizing the Efficacy of Thermal Camera-based Attacks. In *Proceedings of the 5th USENIX Conference on Offensive Technologies (WOOT'11)*. USENIX Association, Berkeley, CA, USA, 6–6. http://dl.acm.org/citation.cfm?id=2028052.2028058
25. Shwetak N. Patel, Jeffrey S. Pierce, and Gregory D. Abowd. 2004. A Gesture-based Authentication Scheme for Untrusted Public Terminals. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, New York, NY, USA, 157–160. DOI: http://dx.doi.org/10.1145/1029632.1029658
26. Trevor Pering, Murali Sundar, John Light, and Roy Want. 2003. Photographic Authentication Through Untrusted Terminals. *IEEE Pervasive Computing* 2, 1 (Jan. 2003), 30–36. DOI: http://dx.doi.org/10.1109/MPRV.2003.1186723
27. Alexander P. Pons and Peter Polak. 2008. Understanding User Perspectives on Biometric Technology. *Commun. ACM* 51, 9 (Sept. 2008), 115–118. DOI: http://dx.doi.org/10.1145/1378727.1389971
28. Florian Schaub, Peter Lang, Bastian Könings, and Michael Weber. 2013. PriCal: Dynamic Privacy Adaptation of Collaborative Calendar Displays. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp '13 Adjunct)*. ACM, New York, NY, USA, 223–226. DOI: http://dx.doi.org/10.1145/2494091.2494163
29. Stefan Schneegass, Frank Steimle, Andreas Bulling, Florian Alt, and Albrecht Schmidt. 2014. SmudgeSafe: Geometric Image Transformations for Smudge-resistant User Authentication. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 775–786. DOI: http://dx.doi.org/10.1145/2632048.2636090
30. Julian Seifert, Alexander De Luca, and Enrico Rukzio. 2012. Don't Queue Up!: User Attitudes Towards Mobile Interactions with Public Terminals. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM '12)*. ACM, New York, NY, USA, Article 45, 4 pages. DOI: http://dx.doi.org/10.1145/2406367.2406422
31. Richard Sharp, James Scott, and Alastair R. Beresford. 2006. *Secure Mobile Computing Via Public Terminals*. Springer Berlin Heidelberg, Berlin, Heidelberg, 238–253. DOI:http://dx.doi.org/10.1007/11748625_15
32. Garth B. D. Shoemaker and Kori M. Inkpen. 2001. Single Display Privacyware: Augmenting Public Displays with Private Information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. ACM, New York, NY, USA, 522–529. DOI: http://dx.doi.org/10.1145/365024.365349
33. Chen Song, Aosen Wang, Kui Ren, and Wenyao Xu. 2016. "EyeVeri: A Secure and Usable Approach for Smartphone User Authentication". In *IEEE International Conference on Computer Communication (INFOCOM'16)*. San Francisco, California, 1 – 9.
34. Martin Stokkenes, Raghavendra Ramachandra, and Christoph Busch. 2016. Biometric Authentication Protocols on Smartphones: An Overview. In *Proceedings of the 9th International Conference on Security of Information and Networks (SIN '16)*. ACM, New York, NY, USA, 136–140. DOI: http://dx.doi.org/10.1145/2947626.2951962

35. Paul Viola and MichaelJ. Jones. 2004. Robust Real-Time Face Detection. *International Journal of Computer Vision* 57, 2 (2004), 137–154. DOI: http://dx.doi.org/10.1023/B:VISI.0000013087.49260.fb

36. Daniel Vogel and Ravin Balakrishnan. 2004. Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, New York, NY, USA, 137–146. DOI: http://dx.doi.org/10.1145/1029632.1029656

37. Emanuel von Zezschwitz, Alexander De Luca, Bruno Brunkow, and Heinrich Hussmann. 2015. SwiPIN: Fast and Secure PIN-Entry on Smartphones. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1403–1406. DOI: http://dx.doi.org/10.1145/2702123.2702212

38. Emanuel von Zezschwitz, Alexander De Luca, and Heinrich Hussmann. 2013. *Survival of the Shortest: A Retrospective Analysis of Influencing Factors on Password Composition*. Springer Berlin Heidelberg, Berlin, Heidelberg, 460–467. DOI: http://dx.doi.org/10.1007/978-3-642-40477-1_28

39. Emanuel von Zezschwitz, Paul Dunphy, and Alexander De Luca. 2013a. Patterns in the Wild: A Field Study of the Usability of Pattern and Pin-based Authentication on Mobile Devices. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '13)*. ACM, New York, NY, USA, 261–270. DOI: http://dx.doi.org/10.1145/2493190.2493231

40. Emanuel von Zezschwitz, Anton Koslow, Alexander De Luca, and Heinrich Hussmann. 2013b. Making Graphic-based Authentication Secure Against Smudge Attacks. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces (IUI '13)*. ACM, New York, NY, USA, 277–286. DOI: http://dx.doi.org/10.1145/2449396.2449432

41. Yanxia Zhang, Andreas Bulling, and Hans Gellersen. 2014. Pupil-canthi-ratio: A Calibration-free Method for Tracking Horizontal Gaze Direction. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces (AVI '14)*. ACM, New York, NY, USA, 129–132. DOI: http://dx.doi.org/10.1145/2598153.2598186

42. Yulong Zhang, Zhaonfeng Chen, Hui Xue, and Tao Wei. 2015a. Fingerprints On Mobile Devices: Abusing and leaking. In *Black Hat Conference*.

43. Yanxia Zhang, Ming Ki Chong, Jörg Müller, Andreas Bulling, and Hans Gellersen. 2015b. Eye tracking for public displays in the wild. *Personal and Ubiquitous Computing* 19, 5 (2015), 967–981. DOI: http://dx.doi.org/10.1007/s00779-015-0866-8